Measuring Selection when Parents and Offspring Interact

Measuring Selection when Parents and Offspring Interact

Caroline E. Thomson
Department of Zoology, Edward Grey Institute, University of Oxford, Oxford, OX1 3PS, United Kingdom
Evolution Diversité Biologique, Bâtiment 4R1, Université de Toulouse Paul Sabatier, 118 Route de Narbonne, 31062 Toulouse Cedex 09, France
caroline.thomson@univ-tlse3.fr

Jarrod D. Hadfield
Institute of Evolutionary Biology, University of Edinburgh, UK
j.hadfield@ed.ac.uk

1 Non-social and social selection gradients are key evolutionary parameters in systems where individuals interact. They are most easily obtained by regressing an individual's fitness on the trait values of the individual and its social partner.

2 In the context of parental care it is more common to regress the trait value of the parents (ie. the social partner) on a 'mixed' fitness measure that is a function of the parent's and offspring's fitness (for example the number of recruits, which equals parental fecundity multiplied by offspring survival).

3 For such an approach to yield correct estimates of net-selection, the trait must be sex-limited and not affect the parents' own survival.

4 When a trait is not sex-limited, the non-social selection should be weighted by one (because all individuals express the trait) and social selection should be weighted by a half (because the relatedness between parents and the offspring they care for is a half, usually). The 'mixed' fitness approach does not give estimates of both components of selection and so they cannot be weighted appropriately.

5 We show that mixed fitness components are frequently used in place of direct fitness measures in the literature (37% of fecundity selection estimates use a mixed fitness approach), but that the frequency is much higher in some taxa, such as birds and mammals.

6 We suggest alternative methods that could be used to estimate both social and non-social selection gradients, while at the same time assessing the importance of unmeasured traits.

When measuring evolutionary change, the time points between which change is measured must be stipulated. In species with discrete generations the most natural time point is at conception such that evolutionary change is due to selection *within* a generation followed by the inheritance of that change *across* generations (Falconer, 1983). In species with overlapping generations all individuals are not conceived at a single point in time, but evolutionary change can still be measured as the difference in breeding value of newly conceived individuals born one unit of time apart (Hill, 1974; Charlesworth, 1994).

In many taxa, parents directly affect the attributes of their offspring either through the properties of their eggs/seeds or through extended post-natal care (Mousseau & Fox, 1998; Clutton-Brock, 1991; Royle *et al.*, 2012). These non-genetic cross-generational effects complicate the study of natural selection and evolutionary change, but Kirkpatrick & Lande (1989) developed a body of theory by which they could be understood. They took a direct fitness approach whereby the fitness of an individual is measured from its conception

and is defined as the number of zygotes it produces. However, the fitness of the individual can depend on its parents, either because parental phenotype has a direct effect on the individual's fitness, or indirectly because the fitness of an individual depends on its own phenotype, which is partly determined be parental phenotype. This model by Kirkpatrick & Lande (1989) (henceforth the K-L model) is most easily understood when selection on the traits is weak and constant, rather than strong and fluctuating. Then, the change in breeding values between newly conceived individuals of successive generations is:

$$\Delta \mathbf{a}^{(I)} = COV(\mathbf{a}^{(I)}, \mathbf{z}^{\top(I)})(\boldsymbol{\delta} \circ \boldsymbol{\beta}^{(I)}) + COV(\mathbf{a}^{(I)}, \mathbf{z}^{\top(S)})\boldsymbol{\beta}^{(S)} \tag{1}$$

where $\mathbf{a}^{(I)}$ and $\mathbf{z}^{(I)}$ are the vectors of breeding values and phenotypes, respectively, in an individual and $\mathbf{z}^{(S)}$ is the vector of phenotypes in that individual's mother (the social partner of the focal offspring). The $^\top$ superscript denotes the vector transpose. $\boldsymbol{\beta}^{(I)}$ is the *direct* effect of the individual's own traits on the individual's fitness, and this is multiplied element wise by $\boldsymbol{\delta}$ (as indicated by $\circ$) which has elements equal to a half if the trait is sex-limited (Lande, 1982) and one otherwise. We refer to $\boldsymbol{\beta}^{(I)}$ throughout as the non-social selection gradient (Wolf *et al.*, 1999), although it has also been called a direct selection gradient (Kirkpatrick & Lande, 1989; Hadfield, 2012). $\boldsymbol{\beta}^{(S)}$ is the *direct* effect of the individual's mother's traits on the individual's fitness. It is not multiplied by a half because all individuals, both male and female, have a mother, and we assume that the maternal effect is not sex-specific. We call $\boldsymbol{\beta}^{(S)}$ the social selection gradient (Wolf *et al.*, 1999), although it has also been called a parental selection gradient (Kirkpatrick & Lande, 1989; Hadfield, 2012). The first term in Equation 1 can be thought of as the correlated response of breeding values to selection on the individual's own traits, and the second term as the correlated response of breeding values to selection on the individual's parent's traits. The covariance between breeding value and phenotype is complicated when mothers

and offspring interact because maternal effects contribute to the covariance in addition to the direct effect of (inherited) genes. However, as with standard patterns of inheritance $COV(\mathbf{a}^{(I)}, \mathbf{z}^{\top(S)}) = \frac{1}{2}COV(\mathbf{a}^{(I)}, \mathbf{z}^{\top(I)})$ and so Equation 1 is often expressed as

$$\Delta\mathbf{a}^{(I)} = COV(\mathbf{a}^{(I)}, \mathbf{z}^{\top(I)}) \left(\boldsymbol{\delta} \circ \boldsymbol{\beta}^{(I)} + \frac{1}{2}\boldsymbol{\beta}^{(S)} \right) \tag{2}$$

where the selection term in brackets is called the net selection gradient. It should be emphasised, however, that the factor of half associated with social selection is due to inheritance: a different value would be used if mothers were not related to the individuals they care for by half (for example if there was egg dumping (Andersson *et al.*, 2017), or extra-pair paternity in paternal/biparental care models (Thomson *et al.*, 2017)).

Although the K-L model employs a direct fitness approach (where fitness is measured from conception as the number of zygotes produced), it is possible - and useful - to interpret it from an inclusive-fitness perspective (Hadfield & Thomson, 2017). Indeed, the two approaches yield the same results, but from a mathematical perspective the direct fitness approach is often simpler when constructing theoretical models (Taylor *et al.*, 2007) and, we argue, when applying statistical models to data. However, in many empirical studies of natural selection, the number of *recruits* an individual leaves is often advocated as a fitness measure (Clutton-Brock, 1988; Moran & Clark, 2012), which we call a 'mixed' fitness measure, as it combines both parental fitness (fecundity) and offspring fitness (survival). This is neither a direct fitness nor inclusive fitness approach (Grafen, 1982), and the resulting selection estimates have no easy evolutionary interpretation. Because of this, there have been repeated calls, primarily from evolutionary geneticists, to measure fitness from conception (Arnold, 1985; Cheverud & Moore, 1994; Hadfield, 2012; Smiseth *et al.*, 2012). In contrast, the most thorough theoretical work exploring the consequences of using a mixed

fitness measure seems to suggest that both the direct fitness approach and the mixed fitness approach have shortcomings (Wolf & Wade, 2001). Here we reappraise the value of the mixed fitness approach and show that, in general, it will give the wrong answer. The conditions under which it gives the right answer are quite restrictive, in contrast to the direct fitness approach that, if used correctly, can be applied in a wide range of circumstances.

As in Wolf & Wade (2001), our immediate criterion for correctness is whether the estimated selection gradient multiplied by the genetic variance correctly predicts the amount of evolutionary change. However, this is probably a secondary aim of most biologists, who are often more interested in quantifying selection to understand the adaptive significance of the traits they study (Grafen, 1988). In these instances the 'mixed' fitness approach usually obscures the underlying biology by conflating inheritance and selection, and the fitness of parents and their offspring. This conflation prevents the clean assessment of patterns of natural selection and makes the study of ideas such as parent-offspring conflict exceptionally difficult (Smiseth *et al.*, 2012; Hadfield, 2012).

Theory

The most general model in Wolf & Wade (2001) follows that of Cheverud's (1984) extension of the Willham (1972) model. Two traits are considered where trait 1 maternally affects trait 2, with maternal effect coefficient $\psi_{2,1}$. Non-social selection acts on both traits, but social selection only acts on trait 1. Social selection on trait 1 is assumed to affect fitness through juvenile survival only, and Wolf & Wade (2001) also assume that non-social selection on trait 1 is limited to fecundity, and non-social selection on trait 2 is limited to juvenile survival. Here we relax these assumptions and allow non-social selection on both traits to operate through both fitness components. The two fitness components are given by:

$$w^{(I:J)} = \mu^{(J)} + \beta_1^{(I:J)} z_1^{(I)} + \beta_1^{(S:J)} z_1^{(S)} + \beta_2^{(I:J)} z_2^{(I)} \tag{3}$$

and

$$w^{(I:F)} = \mu^{(F)} + \beta_1^{(I:F)} z_1^{(I)} + \beta_2^{(I:F)} z_2^{(I)} \tag{4}$$

where w is realtive fitness and we use the notation $:J$ or $:F$ to denote quantities that relate to juvenile survival and adult fecundity respectively, and I: or S: to indicate that the trait is expressed in the individual or its social partner, respectively. Assuming our organisms are semelparous total absolute fitness $W^{(I)}$ is simply $W^{(I:J)}W^{(I:F)}$. In Figure 1 a graph of the causal relationships between traits, and traits and fitness components is given.

Figure 1 here

However, in many studies the fitness measure is not the survival and fecundity of a single individual, but often the fecundity of an individual multiplied by the survival of that individual's offspring (e.g. number of recruits): $W^{(M)} = W^{(I:J)}W^{(S:F)}$ where the superscript M stands for mixed. Arnold & Wade (1984a,b) show that when selection is weak and the total lifetime fitness of individuals can be divided into multiplicative episodes, then selection gradients can obtained by regressing the relative fitness at each episode on trait values, and then summing the gradients across episodes. Wolf & Wade (2001) consider two approaches for obtaining an estimated selection gradient for trait 1: *Direct*: the univariate regression of direct relative fitness ($w^{(I)}$) on trait 1 of the individual ($z_1^{(I)}$) and *Mixed*: the univariate regression of the relative number of recruits ($w^{(M)}$) on trait 1 of the parent ($z_1^{(S)}$). In what follows we will also deviate from Wolf & Wade (2001) and relax the assumption that trait 1 has to be sex-limited and allow environmental covariances between the two traits as well as genetic covariances.

Assuming that trait 1 is variance standardised, the estimated selection gradient using mixed fitness is (Robertson, 1966; Lande & Arnold, 1983; Arnold & Wade, 1984a):

$$\begin{aligned}\beta_{1,M} = &\quad \delta COV(w^{(I:J)} + w^{(S:F)}, z_1^{(S)}) \\ = &\quad \delta \tfrac{1}{2}\beta_1^{(I:J)} g_1 + \delta\beta_1^{(S:J)} + \delta(\beta_2^{(I:J)} + \beta_2^{(I:F)})(\tfrac{1}{2}g_{1,2} + \psi_{2,1}) + \delta\beta_1^{(I:F)}\end{aligned} \tag{5}$$

which is equivalent to Equation A8 in Wolf & Wade (2001) if $\beta_1^{(I:J)} = \beta_2^{(I:F)} = 0$ and $\delta = 1/2$. In the results section we consider a number of special cases of this general equation and discuss how, or even if, the resulting mixed selection gradients can be interpreted.

Using direct fitness in the univariate approach the estimated selection gradient is

$$\begin{aligned}\beta_{1,D} = &\quad COV(w^{(I:J)} + w^{(I:F)}, z_1^{(I)}) \\ = &\quad \beta_1^{(I:J)} + \tfrac{1}{2}\beta_1^{(S:J)} g_1 + (\beta_2^{(I:J)} + \beta_2^{(I:F)})(g_{1,2} + e_{1,2} + \tfrac{1}{2}\psi_{2,1} g_1)\end{aligned} \tag{6}$$

which is Equation A7 in Wolf & Wade (2001) when $\beta_1^{(I:J)} = \beta_2^{(I:F)} = 0$, $e_{1,2} = 0$ and $\delta = 1/2$. However, although direct fitness is used, the method does not include both offspring and parental traits as predictors of an individual's fitness and therefore is not appropriate for estimating social selection gradients (Kirkpatrick & Lande, 1989). Because of this, we do not discuss this approach further, and instead spend time discussing simple methods for estimating social and non-social selection gradients using a direct fitness approach.

The measure of validity used by Wolf & Wade (2001) was whether the estimated selection gradients multiplied by the genetic variance will correctly predict the amount of evolutionary change in trait 1. From Equation 2 we can see that this will be the case when the estimated selection gradient is equal to what Kirkpatrick & Lande (1989) call the net selection gradient:

$$\beta_{1,M} = \delta_1 \beta_1^{(I)} + \frac{1}{2}\beta_1^{(S)} \tag{7}$$

and the covariance between the breeding value for trait 1 and the phenotype for trait 2 is zero (or net selection on trait 2 is zero):

$$\mathrm{COV}(a_1^{(I)}, z_2^{(I)}) = \mathrm{COV}(a_1^{(I)}, z_2^{(S)}) = 0 \quad \text{or} \quad \delta_2 \beta_2^{(I)} + \frac{1}{2}\beta_2^{(S)} = 0 \tag{8}$$

In cases where relatedness between parents and the offspring they care for is not a half, the factor of a half can be replaced by relatedness in the preceding (and following) equations.

To evaluate when these two conditions will be met, and why, we work through a series of examples that have different patterns of selection and maternal effects. In all cases, we assume that selection has been measured through the effect of trait 1 (z_1) on a mixed fitness measure (number of offspring surviving to some point past conception). First, we consider the case where trait 2 is absent, and trait 1 alone affects the individual's fecundity and the individual's offspring's survival (Figure 2); the case which the mixed fitness approach seems to be most suited to. Then we consider more complicated scenarios where the mixed fitness approach would appear less suited (Figure 3).

If the causal model motivating the mixed fitness approach is true

Figure 2 here

In the first instance, we will assume that the causal model that appears to motivate the mixed fitness approach is true (Figure 2): trait 1 can affect its bearer's own fitness via fecundity and that of its offspring via survival. All other routes by which trait 1 could affect the fitness of either party are assumed absent, and a second (unmeasured) trait is assumed not to exist. In this case

fitness via survival is simply $w^{(I:J)} = \mu^{(J)} + \beta_1^{(S:J)} z_1^S$ and the mixed selection gradient is therefore:

$$\beta_{1,M} = \delta_1 \beta_1^{(I:F)} + \delta_1 \beta_1^{(S:J)} \tag{9}$$

i.e. the sum of non-social selection through the effects of trait 1 on fecundity ($\beta_1^{(I:F)}$), and social selection through effects of the social partner's trait 1 on survival ($\beta_1^{(S:J)}$). Under these conditions Wolf & Wade (2001) state that the mixed fitness approach is a valid way of estimating net-selection, but we see here that this relies on the assumption that the trait is sex-limited, i.e. $\delta_1 = \frac{1}{2}$ (as acknowledged by Wolf & Wade, 2001) and that the relatedness of parents and the offspring they care for is a half. In addition the mixed fitness approach does not allow the researcher to get individual estimates of social and non-social selection. However, if the trait does not affect the parent's own fecundity then $\beta_{1,M} = \delta_1 \beta_1^{(S:J)}$ is a valid social selection gradient (although halved if it is assumed trait 1 is sex-limited). Alternatively, if the trait does not affect the offspring's survival then $\beta_{1,M} = \delta_1 \beta_1^{(I:F)}$ and is a valid non-social selection gradient.

If the causal model motivating the mixed fitness approach is not true

The case presented above assumes that the underlying model is that for which the mixed fitness approach is most suited. However, one can envision many situations where the biology is more complicated (Figure 3). Below we add additional fitness and maternal effects to the basic model described above, and illustrate the model with a *possible* example from the literature. Similar scenarios to ii) and iii) are also covered in Wolf & Wade (2001) with sex-limitation.

Figure 3 here

(i) Juvenile survival is affected by the individual's own trait 1 ($\beta_1^{(I:J)} \neq 0$).

In this case, the trait is simultaneously expressed in both parents and their offspring. A possible example of such a scenario is provided by Bouteiller-Reuter & Perrin (2005) who estimated selection on female body mass in greater white-toothed shrews (*Crocidura russula*) using the number of weaned offspring per litter as a fitness measure. Body mass is evidently expressed in both parents and their offspring simultaneously, and so juvenile survival may be influenced by the individual's own trait value rather than (or in addition to) that of the parent. Consequently, juvenile survival is determined by both its own trait value ($z_1^{(I)}$) and that of the mother ($z_1^{(S)}$), and becomes $w^{(I:J)} = \mu^{(J)} + \beta_1^{(I:J)} z_1^{(I)} + \beta_1^{(S:J)} z_1^{(S)}$ and

$$\beta_{1,M} = \delta_1 \beta_1^{(I:F)} + \delta_1 \beta_1^{(S:J)} + \delta_1 \frac{1}{2} g_1 \beta_1^{(I:J)} \tag{10}$$

Where, as before, $\beta_1^{(I:J)}$ is the non-social selection acting through juvenile survival, and $\beta_1^{(S:J)}$ is the social selection gradient acting through juvenile survival. Here the mixed fitness approach cannot give the correct answer even when the trait is sex-limited, because direct (non-social) selection operating through juvenile survival is underestimated by a factor equal to half the heritability ($g_1 = h_1^2$ because the trait has been variance standardised). Similarly, in cases where parental traits have no direct effect on offspring fitness, after conditioning on offspring traits (there is no social selection; $\beta_1^{(S:J)} = 0$) then:

$$\beta_{1,M} = \delta_1 \beta_1^{(I:F)} + \delta_1 \frac{1}{2} g_1 \beta_1^{(I:J)} \tag{11}$$

If non-social selection on a trait via fecundity was antagonistic to that on juvenile survival (i.e. $\beta_1^{(I:F)}$ and $\beta_1^{(I:J)}$ have opposing signs), such that there was no overall effect of the trait on fitness, then using a mixed fitness measure would incorrectly provide evidence of selection for trait values that favour fecundity.

(ii) Fitness is affected by a second (non-sex-limited) trait expressed in the

individual ($\beta_2^{(I:J)} + \beta_2^{(I:F)} \neq 0$).

Whilst the preceding considerations are likely to be important when selection is measured on a trait that is expressed concurrently in two generations (such as body size), many studies estimate selection on traits only expressed in adults. For example, selection on phenological traits has been estimated through their effects on offspring fitness, including the effects of laying date (Charmantier *et al.*, 2006), parturition date (McAdam & Boutin, 2003), and arrival date (Seamons *et al.*, 2007). As the trait is not expressed during juvenile life stages $\beta_1^{(I:J)} = 0$ by definition. However, a second trait (z_2) expressed at juvenile, and possibly adult stages (such as body mass), may be genetically correlated with the focal trait ($g_{1,2} \neq 0$). This is illustrated well by Sheldon *et al.* (2003), where a significant genetic correlation exists between laying date and tail length in collared flycatchers, and both are shown to be under significant directional selection through a measure of mixed fitness. In such cases, where trait 1 is only expressed in adults but is genetically correlated with trait 2, then selection measured on trait 1 becomes

$$\beta_{1,M} = \delta_1\beta_1^{(I:F)} + \delta_1\beta_1^{(S:J)} + \delta_1(\beta_2^{(I:J)} + \beta_2^{(I:F)})\frac{1}{2}g_{1,2} \tag{12}$$

Where $\beta_2^{(I:J)}$ and $\beta_2^{(I:F)}$ are the non-social selection gradients on the second trait acting through effects on juvenile survival and adult fecundity, respectively. Only when the genetic correlation between the traits is zero ($g_{1,2} = 0$), and trait 1 is sex-limited ($\delta_1 = \frac{1}{2}$), does the mixed fitness approach give the correct answer. More generally, selection on trait 1 will be biased towards the correlated response to selection on trait 2 ($g_{1,2}\beta_2^{(I)}$) although this will be multiplied by a half, or a quarter if trait 1 is assumed to be sex-limited (see Cheverud, 1984, also).

(iii) Fitness is affected by a second (non-sex-limited) trait expressed in the

individual that is maternally affected by trait 1 ($\psi_{2,1} \neq 0$ and $\beta_2^{(I:J)} + \beta_2^{(I:F)} \neq 0$). In the above example, trait 1 in the mother and trait 2 in the offspring are assumed to be correlated because of shared genes. However, a correlation may also exist if trait 1 in the parents directly affects the value of the second trait in the offspring, through maternal (or paternal) effects. For example, in Thomson *et al.* (2017) we show how parental performance for offspring mass (a trait of the parents that captures all effects they have on their offspring's mass) directly affects the parent's own fecundity and indirectly affects their offspring's survival via an effect on body mass. Under scenarios like this,

$$\beta_{1,M} = \delta_1\beta_1^{(I:F)} + \delta_1\beta_1^{(S:J)} + \delta_1(\beta_2^{(I:J)} + \beta_2^{(I:F)})(\tfrac{1}{2}g_{1,2} + \psi_{2,1}) \quad (13)$$

which is equivalent to scenario ii) but the term $\psi_{2,1}$ (the maternal effect coefficient) contributes to the covariance between the traits. As a consequence, the mixed fitness approach fails when the trait maternally affects other traits under selection even when the traits are not genetically correlated.

Empirical Patterns

Methods

To assess the frequency with which direct and mixed fitness approaches are used in the literature, we went through the papers from which Kingsolver & Diamond (2011) had collated estimates of selection from wild populations. For each paper we assessed which of the two approaches the authors had used when estimating linear selection gradients. These papers are a combination of those from Kingsolver *et al.* (2001) and Siepielski *et al.* (2009), and inclusion criteria are explained explicitly in those papers. Broadly, the papers report all studies of selection on quantitative phenotypic traits from wild unmanipulated populations from 1984 to 2001 (Kingsolver *et al.*, 2001), and all studies with temporally

replicated estimates from 2002 to 2009 (Siepielski *et al.*, 2009). In total there are 2819 estimates of linear selection gradients from 97 studies across 89 species. The fitness measures used in these studies had already been broadly categorised as fecundity, mating success, survival, and total (lifetime) selection, but there was often heterogeneity within a category with regard to which fitness components had actually been measured, and whether the fitness components were measured on the same individual or different individuals. In particular, the original 'fecundity' category often included measures of the number of zygotes/eggs, but also the number of surviving offspring. Thus, we explicitly recategorised the measures used in these studies as Adult survival (A), Juvenile Survival (J; any survival pre-recruitment was considered juvenile), Mating Success (M), and Fecundity (F; the number of zygotes) and studies using fitness measures that were a composite were recorded as such. In addition, for those studies where the fitness measure could not be truly classified into one of these categories, we recorded it as 'other' (using a proxy for fitness, or the true measure could not be determined from the paper). We also recorded whether the trait was measured in the same individual for which the fitness component was defined, or on the individual's parent. For example, a study that looked at selection on a parental trait where the fitness measure was how many offspring that individual recruited into the population would be denoted as $F^{(S,S)} + J^{(I,S)}$ where the first letter in the superscript designates whose fitness was measured and the second letter whose trait was measured.

Results

Of the 2819 estimated linear selection gradients in Kingsolver & Diamond (2011), the fitness measure used could be classified according to our system for 2556 estimates from 95 papers. Reclassification of the fitness measures showed that there was considerable disparity between studies classified under the same

original broad fitness measure categories, and in total 637 selection gradients from 22 studies used either a mixed fitness approach or had measured social selection gradients rather than non-social selection gradients. Originally, 681 selection gradients had been classed as measuring selection using fecundity as a fitness measure. From our reclassification, 351 truly measured fecundity as the number of zygotes, with a further 66 as some measure of mating success (generally pollen grains removed). 252 selection gradients were estimated using a mixed fitness approach (106 of which considered offspring survival pre-independence, and the others post-independence from the parent). Of the 602 selection gradients classed as measuring fitness as mating success, 185 used a mixed fitness approach (of which 84 included pre-independence offspring survival, and and the other 101 used a post-independence time point).

Furthermore, there are 1263 measures of survival selection, of which 94 used mixed fitness. In addition, 74 social selection gradients were measured - 28 used a measure of survival to independence, 36 used survival to a point post-independence, and 10 used offspring survival from a point post-independence to another time point (recruitment, or pupation). Finally, of the 52 selection gradients reporting 'total' fitness, 20 used a mixed fitness approach.

The distribution of the different fitness measures across taxonomic groups can be seen in Table 2. This makes it clear that there are differences between researchers working in different taxonomic domains in how fitness is measured; whilst true fecundity (as the number of zygotes, or some proxy for this) is often measured in Angiosperms and insects, a mixed fitness measure of fecundity is more frequently measured in birds.

Table 2 here

Statistical Solutions

Methods

We propose a statistical method that simultaneously models both survival and fecundity, and allows both non-social and social selection gradients to be estimated. The advantage of modelling survival and fecundity simultaneously, rather than in separate analyses, is that it allows any remaining covariance (after conditioning on measured traits) between the parent's fecundity and the offspring's survival to be estimated. For analysing data from the full model described above we imagine two statistical models:

$$f_i = b_0^{(F)} + z_{1i} b_1^{(I:F)} + z_{2i} b_2^{(I:F)} + e_i^{(F)} \tag{14}$$

where f_i is the linear predictor for the fecundity of individual i, $b_0^{(F)}$ is the intercept, $b_1^{(I:F)}$ the regression coefficient associated with the indivdiuals' own trait 1 values and e_i the residual. In what follows we will assume that the fecundity of an an individual is Poisson distributed with rate $\exp(f_i)$. s_{ij} is the linear predictor for survival of offspring j from individual i:

$$s_{ij} = b_0^{(J)} + z_{1i} b_1^{(S:J)} + z_{1ij} b_1^{(I:J)} + z_{2ij} b_2^{(I:J)} + u_i^{(J)} \tag{15}$$

where $u_i^{(J)}$ is a random effect that allows the survival of offspring from the same parent to be correlated after conditioning on the traits and their associated regression coefficients b. We will assume that the survival probability of individual ij is $F_N(s_{ij})$ where F_N is the Gaussian cumulative density function (i.e a probit or threshold model Pearson, 1900). We allow $e_i^{(F)}$ (the residual parental fecundity) and $u_i^{(J)}$ (the parental effect on offspring survival) to be correlated. Allowing a covariance between the a residual and a random effect is non-standard, but Thomson *et al.* (2017) provides a Markov chain Monte Carlo (MCMC) strategy for estimating such covariances. In this context, the covariance could be due to unmeasured traits that a) have a non-social effect on fecundity and a social effect on survival b) have a non-social effect on fecundity and are phenotypically correlated with other unmeasured traits that have social

effects on survival or c) have a non-social effect on fecundity and survival and are heritable (due to genes and/or maternal effects) or d) have a cross-generation correlation with other unmeasured traits that have a non-social effect on survival. With sufficient data the contribution of a) and b) versus c) and d) could be assessed by including a genetic and/or maternal genetic terms for both fitness components, using cross-fostering or a multigenerational pedigree.

In summary we run a bivariate mixed model with the two responses being fecundity and survival. The individual's own trait values are fitted as fixed effects for each response (non-social selection), and parental trait 1 is included as a fixed effect for juvenile survival (social selection). It should be noted that a social selection effect for trait 2 has not been fitted; it could be fitted, but we prefer to omit a trait so that the code presented in the Supplementary Materials can be more easily tailored to situations where all traits are not fitted as both social and non-social predictors of fitness. This approach for estimating selection is consistent with the K-L approach and has some similarities to contextual analysis (Heisler & Damuth, 1987; Goodnight *et al.*, 1992) and particularly neighbourhood models (Nunney, 1985).

1000 data-sets were simulated according to the model for 200 adults and their offspring. The genetic and environmental variances for z_1 and z_2 were both 1, with genetic covariance $g_{1,2} = 0.25$, random-residual covariance $COV(e^{(F)}, u^{(J)}) = -0.25$ and maternal effect $\psi_{2,1} = 0$. The remaining parameters were all estimated in the model and the values used in the simulation are reported in Table 1. Model parameters were estimated using MCMCglmm (Hadfield, 2010) with flat improper priors, a chain length of 13000, a burn-in of 3000 and a thinning interval of 10. The posterior means and 95% credible intervals for all parameters were stored for each analysis.

The selection gradient is defined as (Lande, 1979; Lande & Arnold, 1983):

$$\boldsymbol{\beta} = \frac{E\left[\partial W/\partial \mathbf{z}\right]}{E\left[W\right]} \tag{16}$$

where W is absolute fitness and $\mathbf{z}$ are all traits, both those of the mother and the individual itself. The expectation is taken over individuals with respect to all variables (in this case the traits and random effects). An element k of the vector of selection gradients ($\boldsymbol{\beta}$) therefore represents the average effect on fitness of perturbing trait k whilst holding the other traits constant. Conditional on the fixed effects and random effects (including the fecundity residual), survival and fecundity are independent such that the expected absolute fitness of an individual is $W_i = F_N(s_i)\exp(f_i)$. Consequently, the k^{th} element of the selection gradient is given as:

$$\boldsymbol{\beta} = \frac{E\left[\exp(f)\left(f_N(s)b_k^{(J)} + F_N(s)b_k^{(F)}\right)\right]}{E\left[F_N(s)\exp(f)\right]} \tag{17}$$

where f_N is the probability density function of the Gaussian. The six dimensional integral required to obtain the expectation is not analyticaly tractable, and so we simply sample the variables from their distribution, evaluate the numerator and denominator in Equation 17 for each sample, and take their averages. If selection is weak, the distribution is

$$\begin{bmatrix} e^{(F)} \\ u^{(J)} \\ \mathbf{z}^{(I)} \\ \mathbf{z}^{(S)} \end{bmatrix} \; N\left(\begin{bmatrix} 0 \\ 0 \\ \boldsymbol{\mu}_{\mathbf{z}} \\ \boldsymbol{\mu}_{\mathbf{z}} \end{bmatrix}, \begin{bmatrix} \mathrm{VAR}(e^{(F)}) & \mathrm{COV}(e^{(F)}, u^{(J)}) & 0 & 0 \\ \mathrm{COV}(u^{(J)}, e^{(F)}) & \mathrm{VAR}(u^{(F)}) & 0 & 0 \\ 0 & 0 & \mathbf{G}+\mathbf{E} & \frac{1}{2}\mathbf{G} \\ 0 & 0 & \frac{1}{2}\mathbf{G} & \mathbf{G}+\mathbf{E} \end{bmatrix}\right) \tag{18}$$

where $\boldsymbol{\mu}_z$ is vector of trait means, and $\mathbf{G}$ and $\mathbf{E}$ are their genetic and environmental covariance matrices. Since the traits are not modelled (they just appear as fixed predictors of survival and/or fecundity) $\boldsymbol{\mu}_z$, $\mathbf{G}$ and $\mathbf{E}$ are not estimated as part of the model. It would be possible to jointly model the distri-

bution of these traits (Thomson *et al.*, 2017) but an easier solution is to assume that the means and (co)variance structure of the four traits (individual and parental) are identical to those actually sampled (i.e. the empirical mean and covariance of the predictors $z_1^{(I)}$, $z_2^{(I)}$, $z_1^{(S)}$ and $z_2^{(S)}$). Such a strategy may also be more robust to strong selection, given it would measure the distribution of the parental traits post-selection, as required (Kirkpatrick & Lande, 1989). In short, we take a draw from Equation 18 and evaluate the two expressions in Equation 17 that lie within the square brackets. We repeat this 1000 times and then take the average of the evaluation for each expression to obtain the selection gradient in Equation 17. This procedure can be repeated for each MCMC iteration to get a posterior distribution for the selection gradient. The code to simulate the data, fit the model and obtain the selection gradients can be found in the supplementary material.

It should be stressed that in this example we do not have the complete life-history for any individual; we have fecundity data from one generation and survival data from the following generation. We therefore have to assume that patterns of fecundity and survival selection are the same in the two generations. With more complete date then this assumption could be relaxed.

Results

The results of the simulation are reported in Table 1. For all parameters, the mean of the posterior means were close to their true values with location terms generally differing by less than ± 0.01. The mean of the posterior mean variances (VAR($e^{(F)}$) and VAR($u^{(J)}$)) were slightly higher than their true values, as would be expected given their skewed distributions, but the means of the posterior modes were closer (0.997 ± 0.005 for VAR($e^{(F)}$) and 1.011 ± 0.008 for VAR($u^{(J)}$)). The covariance between $e^{(F)}$ and $u^{(J)}$ was close to its true value. Coverage seemed reasonable: on average the lower 95% credible interval was above the true parameter value in 26.2/1000 cases and the upper 95% credible interval

was below the true parameter value in 25.7/1000 cases, roughly in-line with the expectation of 25/1000. Under this particular set of parameters the selection gradients are close to the sum of the two (survival and fecundity) regression coefficients ($\beta_1^{(I)} = -0.199$, $\beta_2^{(I)} = 0.298$, $\beta_1^{(S)} = 0.397$ and $\beta_2^{(S)} = 0$) although in general this won't be the case.

Table 1 here

Discussion

In many taxa, parents can affect the survival, and even fecundity, of their offspring. Because of this, phenotypic selection is often measured using the number of recruits an individual leaves as a fitness measure (Clutton-Brock, 1988). This fitness measure, which we call a mixed fitness measure, is a combination of parental fecundity and offspring survival and is generally inappropriate for estimating phenotypic selection. Our conclusion is largely in agreement with Wolf & Wade (2001) although they suggest that when offspring survival is solely a function of parental traits (Grafen's (1988) 'independence of control') the mixed fitness approach can be appropriate. However, we show that this will only be true when the parental trait can be assumed to be sex-limited. Under this condition we do not need to separate the effect of the trait on the individual's own fecundity (non-social selection) from that on the individual's offspring's survival (social selection) because both components are weighted by a half; the first because the trait is only expressed in half the parents (Lande, 1982), and the second because parents are usually related to the offspring they care for by a half (Kirkpatrick & Lande, 1989). When the trait is not sex limited these components have to be weighted by one and half, respectively, and this cannot be done unless the social and non-social selection are estimated separately. In addition, Wolf & Wade (2001) assume that the trait is only expressed at the adult stage and so cannot directly affect the juvenile survival of the individual

itself. If this condition of 'independence of control' is not met, then the contribution of non-social selection acting via juvenile survival is undervalued by half the heritability, and the net selection gradient will be biased towards selection on fecundity.

At face value, the assumption that a trait is sex-limited and is only expressed at the adult stage seems reasonable; other than in birds, uniparental care is the norm, and by definition parental care is only expressed in adults. However, although the parental effect of a trait may only be manifest in adults of one sex, this does not imply that the trait itself needs to be sex-limited and only expressed at maturity (e.g. body-size; Bouteiller-Reuter & Perrin, 2005). For example, of the 19 bird studies in Kingsolver & Diamond (2011) that included selection via juvenile survival, 11 used a mixed fitness approach, 6 of which involved traits that were not sex-limited and 3 involved traits that were expressed at the juvenile stage. Only one study considered the implications of using the number of recruits as a fitness measure (Sheldon *et al.*, 2003) despite more than half using a fitness measure that was inappropriate for the traits studied.

Wolf & Wade (2001) also suggest that the direct fitness approach has its own shortcomings, and because of this the mixed fitness approach still has some utility. However, it is important to realise that the direct fitness approach analysed by Wolf & Wade (2001) is not the appropriate direct fitness approach for the problem at hand; only the trait value of the individual is used to predict fitness. Faced with the option of using a mixed fitness approach, or using a direct fitness approach that ignores the effect of parental traits on offspring fitness, it is only natural that many researchers advocate the former (Clutton-Brock, 1988; Moran & Clark, 2012). However the direct fitness approach used in the K-L model explicitly requires the trait values of both the individual and its mother be used as predictors of an individuals fitness (Kirkpatrick & Lande, 1989; Hadfield, 2012). Doing so is relatively straightforward, and here we suggest a simple statistical model that directly estimates the social and non-

social selection gradients separately, while accounting for the covariance between parental fecundity and offspring survival that is not accounted for by the traits that are the object of the selection analysis.

Throughout, we have presented the problem of mixed fitness in the context of studies that make assumptions about the action of a parental trait on offspring fitness. However, in some cases, mixed fitness may be used without the explicit assumption that the parental trait is directly affecting offspring fitness. For example, even when parents do not explicitly affect their offspring's fitness, a second argument sometimes given for using mixed fitness is that only offspring that survive to breed are able to contribute to the continuation of the focal individual's germ line. This logic has even been extended to suggest that grand-offspring, or more distant descendants, should be counted as fitness (Hunt *et al.*, 2004; Bolund & Lummaa, 2016). However, doing so will exacerbate the problems we highlight because an individual's trait value will be ever more weakly correlated with those of their more distant descendants and so the force of selection will be underestimated. In addition, the fitnesses of relatives will be correlated, even in the absence of genetic variation, as they are calculated from the same numbers, making the interpretation of inheritance difficult. Given these arguments we find mixed fitness approaches that include the survival and fecundity of distant descendants even more hard to justify than the usual two-generational approach.

It has also been suggested that the use of mixed fitness measures in behavioral ecology stems from the fact that behavioral ecologists are more interested in optimality and adaptation than in predicting evolutionary dynamics (Wolf & Wade, 2001). While we agree that the focus of many behavioural ecologists is comparative statics, and that because of this they may be able to ignore the genetic basis of the traits they study (the Phenotypic gambit: Grafen, 1988), we disagree that this focus justifies the use of the mixed fitness approach. For example, evolutionary conflict over parental care traits (Trivers, 1974), due to

antagonistic effects on the fitness of both parents and their offspring (Williams, 1966; Stearns, 1992), are a central topic in behavioral ecology (Davies *et al.*, 2012). In this context the optimal trait value is a compromise between the cost it directly imposes on the parent and the indirect benefits it provides through increased offspring survival (Cheverud, 1984). The net selection gradient will be zero under these circumstances, yet the mixed fitness approach will only provide evidence of this under the restrictive assumptions outlined above. Moreover, even if these assumptions are met, the mixed fitness approach does not allow researchers to quantify the effect of a trait on each component of inclusive fitness. Consequently, when the net selection gradient is zero it would be impossible to determine whether a trait has important but opposing effects on the fitness of parents and their offspring, or simply has no effect on the fitness of either party. The direct fitness approach of the K-L model allows us to say whether the traits are optimal under a broader range of conditions, and also gives us some insight into why they are optimal.

Here we have shown that the common use of 'mixed' fitness approaches to measuring selection are likely to generate misleading results about the strength and direction of selection, and the evolutionary response to that selection. We acknowledge that, in reality, unless the number of offspring can be counted at the point of conception, all fecundity measures are likely to be a mixed fitness to some extent due to early mortality. Nevertheless, the extent to which fitness measures are mixed can be minimised (e.g. count of the number of offspring at birth rather than at the age of ten days). Thus, we suggest that the widespread use of mixed fitness approaches should be replaced by direct fitness approaches unless a valid case can be made that they work for the particular system under study.

Acknowledgements

We thank Jacob Moorad and Per Smiseth for useful discussions regarding this work, and Joel McGlothlin and two anonymous reviewers for their comments. CET was supported by EPSRC, The Clarendon Fund and Magdalen College and JDH by a Royal Society Fellowship.

Data Accessability

Data in this paper is available from Data Dryad doi:10.5061/dryad.k1r87.

Author contributions

CET re-evaluated the selection measures shown in the empirical results; JDH simulated and developed the statistical analysis; CET and JDH developed the theory and wrote the paper.

References

Andersson, M., Taylor, P.D. & Michalakis, Y. (2017) Helping relatives survive and reproduce: Inclusive fitness and reproductive value in brood parasitism. *The American Naturalist*, **189**, 138–152.

Arnold, S.J. & Wade, M.J. (1984a) On the measurement of natural and sexual selection - applications. *Evolution*, **38**, 720–734.

Arnold, S.J. & Wade, M.J. (1984b) On the measurement of natural and sexual selection - theory. *Evolution*, **38**, 709–719.

Arnold, S. (1985) Quantitative genetic models of sexual selection. *Experientia*, **41**, 1296–1310.

Bolund, E. & Lummaa, V. (2016) The effects of resource availability and the demographic transition on the genetic correlation between number of children and grandchildren in humans. *Heredity*, **118**, 186–192.

Bouteiller-Reuter, C. & Perrin, N. (2005) Sex-specific selective pressures on body mass in the greater white-toothed shrew, *Crocidura russula*. *Journal of evolutionary biology*, **18**, 290–300.

Charlesworth, B. (1994) *Evolution in age-structured populations*. Cambridge University Press Cambridge.

Charmantier, A., Perrins, C., McCleery, R.H. & Sheldon, B.C. (2006) Evolutionary response to selection on clutch size in a long-term study of the mute swan. *The American Naturalist*, **167**, 453–465.

Cheverud, J.M. (1984) Evolution by kin selection - a quantitative genetic model illustrated by maternal performance in mice. *Evolution*, **38**, 766–777.

Cheverud, J. & Moore, A. (1994) Quantitative genetics and the role of the environment provided by relatives in behavioral evolution. C.R.B. Boake, ed., *Quantitative Genetic Studies of Behavioral Evolution*, pp. 67–100. University of Chicago Press, Chicago and London.

Clutton-Brock, T.H., ed. (1988) *Reproductive success*. Univeristy of Chicago Press, Chicago.

Clutton-Brock, T.H. (1991) *The Evolution of Parental Care*. Princeton University Press, Princeton, NJ.

Davies, N.B., Krebs, J.R. & West, S.A. (2012) *An introduction to behavioural ecology*. John Wiley & Sons.

Falconer, D.S. (1983) *Introduction to Quantitative genetics*. Longman Group.

Goodnight, C.J., Schwartz, J.M. & Stevens, L. (1992) Contextual analysis of models of group selection, soft selection, hard selection, and the evolution of altruism. *American Naturalist*, **140**, 743–761.

Grafen, A. (1988) On the uses of data on lifetime reproductive success. T.H. Clutton-Brock, ed., *Reproductive success*, pp. 454–471. Univeristy of Chicago Press, Chicago.

Grafen, A. (1982) How not to measure inclusive fitness. *Nature*, **298**, 425–426.

Hadfield, J.D. & Thomson, C.E. (2017) Interpreting selection when individuals interact. *Methods in Ecology & Evolution*, ***submitted***.

Hadfield, J.D. (2010) MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software*, **33**, 1–22.

Hadfield, J. (2012) The quantitative genetic theory of parental effects. N.J. Royle, P.T. Smiseth & M. Kölliker, eds., *The Evolution of Parental Care*, pp. 267–284. Oxford University Press, Oxford, UK.

Heisler, I.L. & Damuth, J. (1987) A method for analyzing selection in hierarchically structured populations. *American Naturalist*, **130**, 582–602.

Hill, W.G. (1974) Prediction and evaluation of response to selection with overlapping generations. *Animal Production*, **18**, 1.

Hunt, J., Bussiere, L.F., Jennions, M.D. & Brooks, R. (2004) What is genetic quality? *Trends in Ecology & Evolution*, **19**, 329–333.

Kingsolver, J.G. & Diamond, S.E. (2011) Phenotypic selection in natural populations: What limits directional selection? *American Naturalist*, **177**, 346–357.

Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E., Hoang, A., Gibert, P. & Beerli, P. (2001) The strength of phenotypic selection in natural populations. *American Naturalist*, **157**, 245–261.

Kirkpatrick, M. & Lande, R. (1989) The evolution of maternal characters. *Evolution*, **43**, 485–503.

Lande, R. (1979) Quantitative genetic analysis of multivariate evolution, applied to the brain:body size allometry. *Evolution*, **33**, 402–416.

Lande, R. (1982) Rapid origin of sexual isolation and character divergence in a cline. *Evolution*, **36**, 213–223.

Lande, R. & Arnold, S.J. (1983) The measurement of selection on correlated characters. *Evolution*, **37**, 1210–1226.

McAdam, A.G. & Boutin, S. (2003) Variation in viability selection among cohorts of juvenile red squirrels (*Tamiasciurus hudsonicus*). *Evolution*, **57**, 1689–1697.

Moran, E.V. & Clark, J.S. (2012) Causes and consequences of unequal seedling production in forest trees: a case study in red oaks. *Ecology*, **93**, 1082–1094.

Mousseau, T.A. & Fox, C.W. (1998) The adaptive significance of maternal effects. *Trends in Ecology & Evolution*, **13**, 403–407.

Nunney, L. (1985) Group selection, altruism, and structured-deme models. *The American Naturalist*, **126**, 212–230.

Pearson, K. (1900) Mathematical contributions to the theory of evolution. VII. on the correlation of characters not quantitatively measurable. *Philosophical Transactions of the Royal Society A*, **195**, 1–47.

Robertson, A. (1966) A mathematical model of culling process in dairy cattle. *Animal Production*, **8**, 95–108.

Royle, N.J., Smiseth, P.T. & Kölliker, M. (2012) *The evolution of parental care*. Oxford University Press.

Seamons, T.R., Bentzen, P. & Quinn, T.P. (2007) DNA parentage analysis reveals inter-annual variation in selection: results from 19 consecutive brood years in steelhead trout. *Evolutionary Ecology Research*, **9**, 409–431.

Sheldon, B.C., Kruuk, L.E.B. & Merila, J. (2003) Natural selection and inheritance of breeding time and clutch size in the collared flycatcher. *Evolution*, **57**, 406–420.

Siepielski, A.M., DiBattista, J.D. & Carlson, S.M. (2009) It's about time: the temporal dynamics of phenotypic selection in the wild. *Ecology Letters*, **12**, 1261–1276.

Smiseth, P.T., Kölliker, M. & Royle, N.J. (2012) What is parental care? N.J. Royle, P.T. Smiseth & M. Kölliker, eds., *The Evolution of Parental Care*, pp. 1–14. Oxford University Press, Oxford, UK.

Stearns, S. (1992) *The Evolution of Life Histories*. Oxford University Press, Oxford, UK.

Taylor, P., Wild, G. & Gardner, A. (2007) Direct fitness or inclusive fitness: how shall we model kin selection? *Journal of evolutionary biology*, **20**, 301–309.

Thomson, C.E., Bayer, F., Farrell, S., Crouch, M., Mittell, E.A., Heap, E. A. Zurita-Cassinello, M. & Hadfield, J.D. (2017) Selection on parental performance opposes selection for larger body size in a wild population of blue tits. *Evolution*, **71**, 716–732.

Trivers, R.L. (1974) Parent-offspring conflict. *American Zoologist*, **14**, 249–264.

Willham, R.L. (1972) The role of maternal effects in animal breeding: III. Biometrical aspects of maternal effects in animals. *Journal of Animal Science*, **35**, 1288–1293.

690 Williams, G.C. (1966) Natural selection, the costs of reproduction, and a refinement of Lack's principle. *American Naturalist*, **100**, 687–690.

692 Wolf, J.B., Brodie, E.D. & Moore, A.J. (1999) Interacting phenotypes and the evolutionary process: II. Selection resulting from social interactions. *Ameri-*
694 *can Naturalist*, **153**, 254–266.

Wolf, J.B. & Wade, M.J. (2001) On the assignment of fitness to parents and
696 offspring: whose fitness is it and when does it matter? *Journal of Evolutionary Biology*, **14**, 347–356.

698 Supporting Information

700 *Simulation and analysis code* : R script for carrying out the simulation and analysis shown in the statistical solution section.

702

Parameter	True Value	Mean Estimate	<l-95%	>u-95%
$b_0^{(F)}$	1.00	0.987 ± 0.003	22	30
$b_0^{(J)}$	0.00	0.000 ± 0.004	22	25
$b_1^{(I:F)}$	-0.10	-0.100 ± 0.002	32	27
$b_2^{(I:F)}$	0.00	0.002 ± 0.002	40	22
$b_1^{(S:J)}$	0.40	0.412 ± 0.003	31	27
$b_1^{(I:J)}$	-0.10	-0.103 ± 0.001	21	20
$b_2^{(I:J)}$	0.30	0.302 ± 0.001	31	22
VAR$(e^{(F)})$	1.00	1.032 ± 0.005	22	32
VAR$(u^{(J)})$	1.00	1.104 ± 0.009	18	30
COV$(e^{(F)}, u^{(J)})$	-0.25	-0.264 ± 0.005	23	22

Table 1: Table of model parameters and their true values used in the simulations. The Mean Estimate is the mean of the posterior means followed by the standard error of the mean. <l-95% and >u-95% are the number of simulations in which the true value is less than the lower 95% credible interval or greater than the upper 95% credible interval, respectively. If the method has good coverage we expect this to be the case in 25 out of the 1000 simulations for each parameter.

	Amphibian	Angiosperm	Arachnid	Bird	Crustacean	Fish	Gymnosperm	Insect	Mammal	Reptile
$J^{(I,I)}$	0	18	0	164	0	110	0	8	33	24
		(4)		(3)		(10)		(2)	(3)	(8)
$A^{(I,I)}$	0	15	0	465	0	55	0	123	40	40
		(6)		(60)		(9)		(37)	(8)	(4)
$F^{(I,I)}$	19	458	0	17	14	0	0	326	0	6
	(7)	(137)		(9)	(4)			(61)		(6)
$F^{(I,I)}M^{(I,I)}+A^{(I,I)}$	0	0	0	0	0	0	0	6	0	0
								(3)		
$J^{(I,I)}+F^{(I,I)}M^{(I,I)}+A^{(I,I)}$	0	0	0	0	0	0	0	20	0	0
								(5)		
$J^{(I,S)}$	0	0	0	32	0	0	12	6	26	10
				(11)			(6)	(1)	(2)	(5)
$F^{(S,S)}+J^{(I,S)}$	0	12	2	429	0	0	0	0	24	12
		(6)	(2)	(32)					(6)	(6)
$F^{(S,S)}M^{(S,S)}+A^{(S,S)}+J^{(I,S)}$	0	0	0	2	0	64	0	0	6	0
				(2)		(4)			(1)	

Table 2: The number of linear selection gradients reported in Kingsolver & Diamond (2011) for each taxonomic group (columns) and fitness measure (rows). The number in brackets is the number of trait/species combinations, such that a study that reports multiple gradients for a trait over time is only counted once. F is fecundity measured as the number of zygotes, M is mating success, A is adult survival, and J is juvenile survival. Superscripts indicate the individual upon whom the fitness and trait have been measured - where the first superscript indicates the fitness and the second the trait. Thus (I,I) indicates the trait and fitness were measured upon the same individual in a direct approach, and (S,S) in the mixed fitness approach, and (I,S) indicates that the fitness and trait were measured in different individuals (the offspring and parent, respectively).

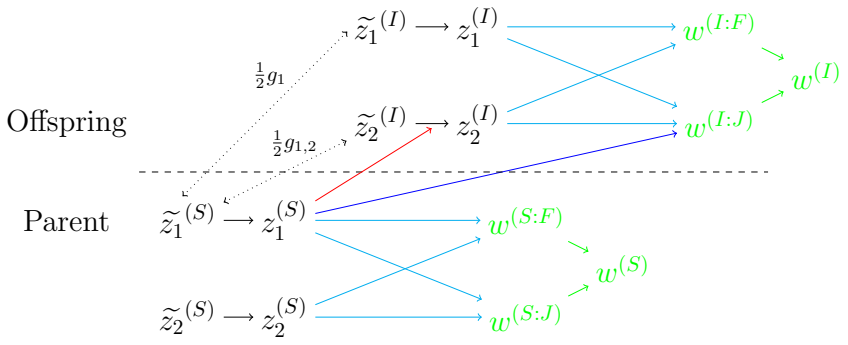

Figure 1: Schematic of the most complex causal model analysed. As in Hadfield & Thomson (2017) we denote the traits prior to the action of maternal effects as $\widetilde{z}_2^{(I)} = a_2^{(I)} + e_2^{(I)}$ and the traits after the action of maternal effects as $z_2^{(I)} = \widetilde{z}_2^{(I)} + \psi_{2,1} z_1^{(S)}$, where a and e are breeding value and environmental value respectively. The red arrow represents the maternal effect of trait 1 on trait 2 and has coefficient $\psi_{2,1}$. Light blue arrows represent non-social selection, and the dark blue arrow represents social selection (on trait 1). The dashed double-headed arrows represent the covariances between the $\widetilde{z}$'s measured in parents and offspring, and are a direct function of the genetic (co)variances. It should be noted that in the presence of maternal effects, the covariance between the z's are not equal to the covariance between the $\widetilde{z}$'s.

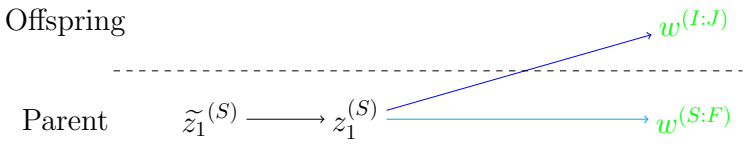

Figure 2: Schematic of the causal model that motivates the use of mixed fitness approach (Equation 9); only trait 1 is considered, which has a direct effect on the fitness of both the parent (S) and the offspring (I) when expressed in the parent. The dark blue arrow represents social selection and the light blue arrow represents non-social selection.

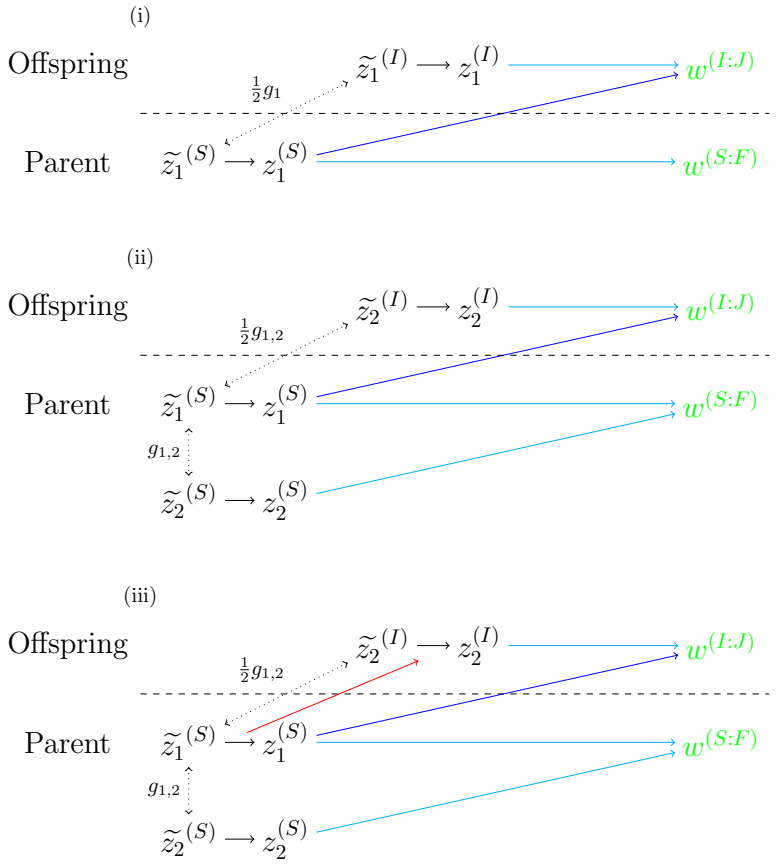

Figure 3: Schematic of the what is measured using a mixed fitness approach when the causal model that motivates its use is not true. (i) the case where trait z_1 affects both the juvenile survival and fecundity of the bearer, and can have a social effect on the juvenile survival of the bearer's offspring (dark blue arrow). (ii) where a second trait (z_2) affects the juvenile survival and fecundity of the bearer and is genetically correlated with z_1. (iii) where a second trait (z_2) affects the juvenile survival and fecundity of the bearer and is maternally affected by z_1.